arXiv:2301.06030v1 [cs.NE] 15 Jan 2023

通过生成式回放实现记忆自我恢复

Zhenglong Zhou、Geshi Yeung 和 Anna C.Schapiro宾夕法尼亚大学心理学系 {zzhou34, geshi, aschapir}@sas.upenn.edu

摘要

大脑一项非凡的能力，是能够在离线阶段自主重组记忆。记忆重放被认为是生物离线学习的基础机制，它启发了人工神经网络在持续学习情境中减少遗忘的离线方法。生成式回放是一种兼顾记忆效率与神经科学合理性的方法，在持续学习基准上达到了最先进的性能。然而，与大脑不同，标准生成式回放在离线阶段使用自身生成的回放样本进行训练时，并不会自我重组记忆。我们提出了一种新颖的架构，为生成式回放增添了适应性强、类似大脑的自主恢复记忆能力。我们在多种持续学习任务和环境中展示了该架构的这一能力。

1 引言

大脑会在离线阶段强化并重组自身的记忆，尤其是在睡眠期间[13, 24, 33]。一种被认为是这种能力基础的机制，是对与清醒体验相关的神经活动进行回放 [5], ；这种现象已在多个物种中得到证实，并被证明有助于记忆 [5, 16]。相关理论认为，在离线阶段重组记忆，是大脑在一生中持续学习新任务、同时不干扰既有知识的关键 [17]。

在人工神经网络（ANNs）中，如何在学习新任务时不遗忘先前任务（即持续学习），仍是一项核心挑战：在非平稳环境下，学习新任务时已无法获取先前任务的数据，融入新信息可能会灾难性地损害已有知识 [18, 6]。受神经科学启发，一种持续学习方法是在学习新任务时回放反映先前任务的数据 [21, 30, 34, 3, 8, 9, 15, 26–29, 25]。以全面反映先前经验的方式回放记忆，可以让来自时间上彼此分隔的经验的信息得到自然整合。最直接的实现方式是将过去任务的真实输入存入记忆缓冲区，之后从中抽取数据进行回放 [26, 28, 3]。然而，在真实世界场景中，回放真实输入难以扩展（例如，通用机器人一生中必须学习大量任务时），而且在生物学上也不合理 [11, 32, 2]。这种方法一种有前景的扩展形式是生成式回放 [30, 34], ：生成模型学习重建先前的输入，从而避免存储精确的输入。这种回放形式被认为在神经层面上更具合理性 [34]；大脑不会存储经验的精确副本，而生物体中的回放类似于从环境生成模型中抽取的带噪样本 [11, 32]。在需要按顺序学习一系列任务的持续学习基准上，生成式回放展现了当前最先进的性能 [35, 34]。

因此，生成式回放似乎是生物学习和人工学习的一种有用机制。然而，尽管生成式回放与神经回放相似，却缺乏类似大脑的能力，无法在离线状态下自适应地自我重组记忆。在这项工作中，我们首先强调了生物回放和人工回放通常评估方式之间的一个关键差异。我们指出，如果采用类似离线处理实证研究的方式进行评估，标准生成式回放并不会改善行为。

第36届神经信息处理系统大会（NeurIPS 2022）上的人工与真实智能研讨会。

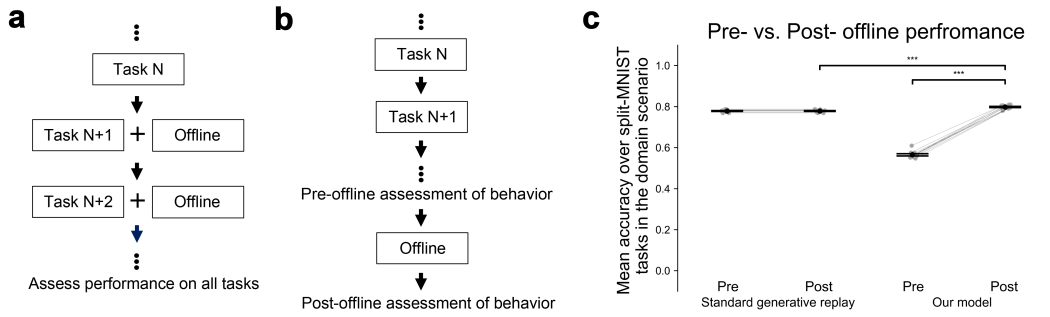

图1：人工与生物离线学习评估方式的差异。**a**。在标准持续学习设置中，离线方法的评估依据是其能否帮助人工神经网络在依次学习一系列任务时保留记忆。**b**。在神经科学和心理学中，离线学习研究考察的是在没有外部输入的情况下，行为如何在离线阶段发生变化。**c**。按照实证研究中的评估方式进行评估时，标准生成式回放并不能改善行为，而我们的模型则显示，离线学习能够提升表现。每个点表示以不同随机种子初始化的模型表现。每个模型实例化前后的离线准确率由灰色线段连接。误差线表示 ± 1 均值标准误。***表示$p<0.001$。

我们提出了一种新颖的架构，使生成式回放具备类似大脑的自我重组能力：通过离线处理自我修复受损记忆的能力。我们在 MNIST [14]和 CIFAR-100 [12] 上展示了该模型具备这一能力。

2 标准生成式回放不会在离线阶段自我重组

心理学和神经科学研究探讨了大脑如何在离线状态下自主重组记忆，从而改善行为。这些研究表明，离线状态下的大脑能够超越清醒状态中所学的内容，通过自发过程引发有益的行为变化 [13, 24, 33]。这些变化通过离线阶段前后的行为差异来衡量（图1b）。例如，学习一系列任务后，一段睡眠可以修复在此前清醒阶段因干扰而受损的记忆 [19, 20, 1, 4] ——生物学细节丰富的睡眠模型已对这一适应能力展开研究[7, 22, 31]。

相比之下，机器学习情境中的标准持续学习评估并未考察离线算法在离线阶段重新组织记忆的能力：在这些情境中，算法的有效性以其帮助人工神经网络在依次学习多个任务时保留过往记忆的效果来衡量（图 1a）。因此，这些算法能否像离线状态下的大脑那样改善行为，尚不明确。值得强调的是，尽管生成式回放在这类情境中能有效减少遗忘，且有人提出它可以模拟生物回放，但模型仅用自身的回放样本训练时，并不会产生任何学习（图 1c）。标准生成式回放 [30] 由两个部分组成：一个学习重建输入的生成器 x，以及一个学习将输入映射到其目标输出的求解器 y。为了回放反映模型对过往任务知识的输入输出对（x'，y'），生成器会采样 x'，以反映其对过往输入学得的重建结果；求解器则根据其学得的映射，将 x' 标记为 y'。将其与新任务的数据混合时，(x'，y') 有助于保留过往任务的知识。然而，如果模型仅用 (x'，y') 训练，其行为不会发生变化，因为 (x'，y') 与求解器学得的输入输出映射完全一致。因此，标准生成式回放无法在离线状态下自行重组记忆。

3 自恢复记忆架构

在本节中，我们提出一种新颖的架构，为生成式回放增添了类似大脑的离线重组记忆能力：自我修复受损记忆的能力。我们使用

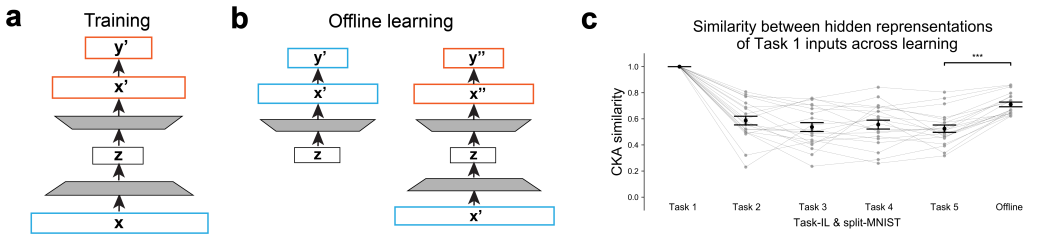

图 2：我们的模型架构及其对隐藏表征的离线恢复。**a**。我们将模型实现为一个带有前馈连接的变分自编码器，这些连接将重建层映射到分类输出。每个阴影梯形表示一组隐藏层。在训练过程中，模型接收输入输出对（x，y），并经过训练输出（x'，y'），以近似（x，y）。在 CIFAR-100 实验中，模型（此处未示出）学习输出（h'，y'），以近似（h，y），其中 h 是 x 的隐藏表征。**b**。在模型完成所有可用任务的训练后，它会通过离线方式从自身生成的回放样本中学习。模型首先生成（x'，y'）对作为离线训练数据（左）。给定 x'，模型输出（x''，y''），并学习最小化（x'，y'）与（x''，y''）之间的差异。在 CIFAR-100 实验中，模型（此处未示出）最小化（h'，y'）与（h''，y''）之间的差异；二者分别是模型的回放样本，以及模型将回放的隐藏表征 h'作为输入时的输出。**c**。对学习阶段中任务 1 输入的隐藏层表征相似性的分析表明，模型能够通过离线回放恢复先前的隐藏表征，甚至包括与最初学习的任务相对应的隐藏表征。该图展示了针对一个隐藏层的此类分析示例——即直接位于潜在层 z 之上的那一层。

两个持续学习基准来评估模型的这一能力：拆分 MNIST [36] 和拆分 CIFAR-100 任务。模型学完所有任务后，我们设置一个离线阶段，让模型完全使用自身生成的回放样本进行学习。为评估自主离线学习，我们测量模型在离线阶段前后于测试数据上的性能变化。

3.1 模型架构

我们提出的模型（图 2a）是一个生成模型，经过训练后能够同时重建每个输入 x 的表征，并生成目标输出 y。我们将该模型实现为一个变分自编码器，通过前馈连接将重建层（即输出输入 x 重建表征的层）映射到一个与 y维度相同、用于分类的输出层。该架构具有神经科学上的合理性：不同于标准生成式回放，该模型不会采用彼此独立的通路来生成输入模式和标签。与先前的工作 [34], 一样，我们训练模型，使其最小化生成损失（即重建损失与变分损失之和）和交叉熵分类损失之和。在离线回放期间，模型首先通过解码器采样 (x'，y') 对。随后，生成的样本 x' 经过模型的完整前向传播，得到模型对x'的重建和分类结果：(x''，y'')。对于离线学习，我们将 (x'，y') 视为目标输出，将 (x''，y'') 视为模型输出（图 2b）。离线模型经过训练，以最小化生成损失与蒸馏损失之和（见 A.2 节）；这两种损失分别针对 (x'，y') 和 (x''，y'') 计算。

3.2 MNIST 上的实验

在第一组实验中，我们在拆分 MNIST任务上测试了模型，该任务将 MNIST数据集划分为五个二分类任务。训练期间，模型依次逐个学习这五个任务。训练时，我们考察了三种持续学习协议，这些协议的区别在于测试时是否提供任务身份（见 A.4 节）：任务增量（即 Task-IL）、域增量（即 Domain-IL）和类别增量学习（即 Class-IL） [35]。学习完成后，模型在离线阶段从自身的回放样本中学习。我们主要关注模型能否在这一离线阶段通过自我学习提升性能。为了留出性能提升空间，在 Task-IL 和 Domain-IL 场景中，我们省略了任务之间的回放，以避免在离线阶段之前就达到性能上限。

表 1：离线准确率变化的后测与前测均值。每项实验中，每个模型使用不同的随机种子运行 10 次，下方显示各次运行准确率的均值（±均值标准误）。

	数据集	任务增量学习	领域增量学习	类别增量学习
我们的模型	拆分 MNIST	3.24 (±0.42)	23.30 (±0.57)	23.91 (±1.28)
BI-R	拆分 MNIST	-1.50 (±1.12)	0.13 (±0.06)	1.55 (±0.22)
标准生成式回放	拆分 MNIST	0.00 (±0.00)	0.00 (±0.00)	0.00（±0.00）
使用打乱目标的对照组	拆分 MNIST	-32.71（±1.17）	-8.68（±1.84）	-26.48（±0.89）
我们的模型	CIFAR-100	3.30 (±0.33)	6.55 (±0.24)	8.59 (±0.38)

为了衡量离线阶段前后性能的变化，我们在离线阶段开始前和结束后，使用同一组留出测试数据对模型进行了测试。在所有场景中，对于以不同随机种子初始化的模型的每次运行，性能都在离线阶段有所提升（表 1）。相比之下，无论是标准生成式回放，还是近期提出的生成式回放改进方法 BI-R， [34] 都未能在离线回放期间实现可靠的性能提升（表 1）。我们通过随机打乱回放数据中 x' 与 y' 的配对进行了对照实验。在这项对照实验中，我们观察到模型在离线阶段的性能有所下降（表 1），这表明 x' 与 y' 的配对对于模型在离线阶段自我恢复记忆至关重要。总之，我们的结果表明，模型能够通过离线回放自我恢复受损的记忆。

3.3 跟踪任务和离线学习中的隐藏表征变化

为了深入了解离线回放如何促进模型恢复记忆，我们在拆分 MNIST 实验中，测量了模型的隐藏表示在不同任务及离线学习期间的变化。为进行这项分析，我们从任务 1 中留出一组输入。我们使用中心核对齐（CKA） [10, 23]，计算这些输入在学习不同阶段的各层表示之间的相似度。我们发现，离线学习有助于恢复各层对任务 1 输入的初始表示（即每层在任务 1 学习结束时的隐藏表示）：在所有层中，初始表示与离线阶段后的表示比与离线阶段前的表示更加相似（三种训练场景中所有层的 CKA 相似度在离线阶段后均较离线阶段前有所提升，$p<0.05$；图 2c 展示了一个示例）。

3.4 将模型扩展至 CIFAR-100

为确定模型的记忆恢复能力能否扩展到更具挑战性的情境，我们在由自然图像类别构成的 CIFAR-100 数据集上测试了模型。我们训练模型依次学习 10 个任务，每个任务都是一个十分类任务，包含 CIFAR-100 中的 10 个图像类别。由于 CIFAR-100 难度较高，在使用该数据集进行模拟时，我们采用了 BI-R——一种生成式回放的改进方法，已被证明在这项任务中能够实现稳健学习 [34]。与我们用于 MNIST 实验的模型一样，我们在重建层之上添加了前馈连接，将重建后的表示映射为分类输出。与 MNIST 实验类似，我们在任务增量学习和领域增量学习场景中禁用了回放，并在所有任务学习结束后，使用回放样本训练模型。在 CIFAR-100 上，我们同样观察到模型在离线阶段期间的表现有所提升。这些结果表明，模型的记忆恢复能力能够泛化到更复杂的任务。

4 结论

在大脑中，离线重放具有生成性，并且似乎能在缺乏外部数据的情况下促进记忆。相比之下，人工系统中的标准生成式回放在使用其内部生成的重放数据进行训练时，并不能提升性能。我们提出一种架构，赋予生成式回放自我重组的能力：自我恢复受损的记忆。该架构是一种生成模型，并通过前馈连接来生成标签。在依次学习拆分 MNIST 和 CIFAR-100 中的一组任务后，我们提出的架构在使用其内部生成的数据进行训练时，能够自我恢复受损的记忆。相比之下，我们没有观察到这种

能力出现在标准生成式回放、近期一种生成式回放扩展中，也没有出现在回放标签被打乱时的我们的模型中。这些结果表明，将任务标签纳入生成过程，或许能使生成式回放模型从现有表征中学习并对其进行自我重组。

参考文献

[1] Bengi Baran、Jessica Wilson 和 Rebecca Spencer。REM依赖性修复竞争性记忆抑制。《实验脑研究》，203(2):471–477，2010年。[2] Margaret F. Carr、Shantanu P. Jadhav 和 Loren M. Frank。清醒状态下的海马回放：记忆巩固与提取的潜在基础。《自然·神经科学》，14(2):147–153，2011年。[3] Arslan Chaudhry、Marcus Rohrbach、Mohamed Elhoseiny、Thalaiyasingam Ajanthan、Puneet K. Dokania、Philip HS Torr 和 Marc'Aurelio Ranzato。持续学习中的微小情景记忆。*arXiv*预印本 *arXiv:1902.10486*，2019。[4] Spyridon Drosopoulos、Claudia Schulze、Stefan Fischer 和 Jan Born。睡眠在记忆自发恢复与巩固中的作用。实验心理学杂志：总论，136(2):169，2007年。[5] David J. Foster。回放日渐成熟。神经科学年度评论，40(581-602):9，2017。[6] Robert M. French。联结主义网络中的灾难性遗忘。认知科学趋势，3(4):128–135，1999年。[7] Oscar C. González、Yury Sokolov、Giri P. Krishnan、Jean Erik Delanois 和 Maxim Bazhenov。睡眠能保护记忆免遭灾难性遗忘吗？*eLife*，9，2020年。[8] Tyler L. Hayes、Giri P. Krishnan、Maxim Bazhenov、Hava T. Siegelmann、Terrence J. Sejnowski 和 Christopher Kanan。深度学习中的回放：当前方法与缺失的生物学要素。神经计算，33(11):2908–2950，2021年。[9] Ronald Kemker 和 Christopher Kanan。Fearnet：用于增量学习的脑启发模型。*arXiv*预印本 *arXiv:1711.10563*，2017。[10] Simon Kornblith、Mohammad Norouzi、Honglak Lee 和 Geoffrey Hinton。重新审视神经网络表征的相似性。载于国际机器学习会议，第3519–3529页。PMLR，2019。[11] Emma L. Krause 和 Jan Drugowitsch。绝大多数清醒海马尖波涟漪都包含具有动量的空间轨迹。《神经元》，110(4):722–733，2022年。[12] Alex Krizhevsky 和 Geoffrey Hinton。从微小图像中学习多层特征。2009年。[13] Nina Landmann、Marion Kuhn、Hannah Piosczyk、Bernd Feige、Chiara Baglioni、Kai Spiegel- halder、Lukas Frase、Dieter Riemann、Annette Sterr 和 Christoph Nissen。睡眠期间的记忆重组。睡眠医学评论，18(6):531–541，2014年。[14] Yann LeCun、Léon Bottou、Yoshua Bengio 和 Patrick Haffner。应用于文档识别的基于梯度的学习。*IEEE*学报，86(11):2278–2324，1998年。[15] Zhizhong Li 和 Derek Hoiem。无遗忘学习。*IEEE*模式分析与机器智能汇刊，40(12):2935–2947，2017。[16] Yunzhe Liu、Matthew M. Nour、Nicolas W. Schuck、Timothy EJ Behrens 和 Raymond J. Dolan。从自发神经活动中解码认知。自然评论：神经科学，23(4):204–214，2022年。

[17] James L. McClelland、Bruce L. McNaughton 和 Randall C. O'Reilly。为何海马体和新皮层中存在互补学习系统：从联结主义学习与记忆模型的成功与失败中获得的启示。《心理学评论》，102(3):419，1995年。[18] Michael McCloskey 和 Neal J. Cohen。联结主义网络中的灾难性干扰：序列学习问题。载于《学习与动机心理学》，第24卷，第109–165页。爱思唯尔，1989年。[19] Elizabeth A. McDevitt、Katherine A. Duggan 和 Sara C. Mednick。快速眼动睡眠可使学习免受干扰。《学习与记忆神经生物学》，122:51–62，2015年。[20] Sara C. Mednick、Ken Nakayama、Jose L. Cantero、Mercedes Atienza、Alicia A. Levin、Neha Pathak 和 Robert Stickgold。小睡对知觉退化的恢复作用。《自然·神经科学》，5(7):677–681，2002年。[21] Volodymyr Mnih、Koray Kavukcuoglu、David Silver、Andrei A. Rusu、Joel Veness、Marc G. Bellemare、Alex Graves、Martin Riedmiller、Andreas K. Fidjeland 和 Georg Ostrovski。通过深度强化学习实现人类水平的控制。《自然》，518(7540):529–533，2015年。[22] Kenneth A. Norman、Ehren L. Newman 和 Adler J. Perotte。减少互补学习系统模型中干扰的方法：振荡抑制与自主记忆复述。《神经网络》，18(9):1212–1228，2005年。[23] Vinay V. Ramasesh、Ethan Dyer 和 Maithra Raghu。灾难性遗忘的剖析：隐藏表征与任务语义。*arXiv*预印本 *arXiv:2007.07400*，2020年。[24] Björn Rasch 和 Jan Born。论睡眠在记忆中的作用。《生理学评论》，2013年。[25] Roger Ratcliff。再认记忆的联结主义模型：学习与遗忘函数施加的约束。《心理学评论》，97(2):285，1990年。[26] Sylvestre-Alvise Rebuffi、Alexander Kolesnikov、Georg Sperl 和 Christoph H. Lampert。iCaRL：增量分类器与表征学习。载于*IEEE*计算机视觉与模式识别会议，第2001–2010页，2017。[27] Anthony Robins。灾难性遗忘、复述与伪复述。《连接科学》，7(2):123–146，1995年。[28] David Rolnick、Arun Ahuja、Jonathan Schwarz、Timothy Lillicrap 和 Gregory Wayne。持续学习中的经验回放。《神经信息处理系统进展》，32，2019。[29] Jonathan Schwarz、Wojciech Czarnecki、Jelena Luketina、Agnieszka Grabska-Barwinska、Yee Whye Teh、Razvan Pascanu 和 Raia Hadsell。Progress & compress：持续学习的可扩展框架。载于国际机器学习会议，第4528–4537页。PMLR，2018年。[30] Hanul Shin、Jung Kwon Lee、Jaehong Kim 和 Jiwon Kim。通过深度生成式回放实现持续学习。《神经信息处理系统进展》，30，2017。[31] Dhairyya Singh、Kenneth A. Norman 和 Anna C. Schapiro。驱动睡眠依赖性记忆巩固的海马体与新皮层自主交互模型。《美国国家科学院院刊》，119(44)，2022年。[32] Federico Stella、Peter Baracskay、Joseph O'Neill 和 Jozsef Csicsvari。海马体重新激活类似布朗扩散的随机轨迹。《神经元》，102(2):450–461，2019。[33] Robert Stickgold。睡眠依赖性记忆巩固。《自然》，437(7063):1272–1278，2005年。[34] Gido M. van de Ven、Hava T. Siegelmann 和 Andreas S. Tolias。用于人工神经网络持续学习的脑启发式回放。《自然·通讯》，11(1):1–14，2020年。

[35] Gido M. van de Ven 和 Andreas S. Tolias。持续学习的三种情景。*arXiv* 预印本 *arXiv:1904.07734*，2019。[36] Friedemann Zenke、Ben Poole 和 Surya Ganguli。通过突触智能实现持续学习。载于国际机器学习会议，第 3987–3995 页。PMLR，2017。

A. 补充信息

A.1 网络架构

我们提出的模型架构是一种带有用于分类的前馈连接的变分自编码器。该模型由以下部分组成：(i) 编码器网络 e_ϕ，将输入向量 x映射为潜变量向量 z，其中每个单元由两个用于参数化高斯分布的值组成（即，每个单元的两个值分别表示高斯分布的均值和标准差）；(ii) 解码器网络 d_ψ，学习将 z 映射为一个向量x'，使其近似于输入 x；以及 (iii) 一个前馈分类器 c_ω，将 x' 映射为用于分类的输出 y'。e_ϕ、d_ψ和 c_ω 的模型参数与 BI-R [34]中报告的参数基本一致。在 MNIST 实验中，e_ϕ 和 d_ψ 均为全连接网络，各由两个隐藏层组成，每个隐藏层包含 400 个隐藏单元，并使用 ReLU 激活函数。在 CIFAR 实验中，e_ϕ 由五个预训练卷积层组成，这些卷积层与 BI-R [34], 所采用的卷积层相同，此外还包含两个全连接层，每层有 2000 个隐藏 ReLU 单元；而 d_ψ 包含两个全连接层，每层有 2000 个隐藏 ReLU 单元。在两项实验中，c_ω、标准生成式回放和 BI-R 的架构均与 van de Ven 等人 [34]所采用的架构相同。

我们提出的模型架构不同于标准生成式回放和 BI-R。在标准生成式回放中，生成模型（即生成器）和分类器之间没有任何连接，并且分别进行训练（即生成模型在图像上训练，分类器在标签上训练）。相比之下，我们的模型将生成器（即我们实现中的变分自编码器）和分类器合并为一个模型。在外部输入 x上训练期间，分类器将 x'——生成器对 x 的重建——作为输入。在我们的模型中，分类器将解码器的输出作为输入，这使任务标签能够成为生成（即解码）过程的一部分。相比之下，在 BI-R 中，分类器将编码器中的隐藏表示作为输入。

A.2 损失函数

模型通过学习来最小化损失 L_{total}:

$$L_{total} = L_r + L_v + L_c$$

其中，L_r 是重建损失，L_v 是变分损失，L_c 是分类损失。

L_v用于衡量编码器输出所指定的高斯分布偏离标准正态分布的程度，对于每个样本**x**，其定义如下:

$$L_\text{v}(\mathbf{x};\phi) = \frac{1}{2}\sum_{i=1}^{D}(1+\log(\sigma_j^{(\mathbf{x})^2}) - \mu_j^{(\mathbf{x})^2} - \sigma_j^{(\mathbf{x})^2})$$

其中，$\mu_j^{(\mathbf{x})}$ 和 $\sigma_j^{(\mathbf{x})}$ 分别是 $\mu^{(\mathbf{x})}$ 和 $\sigma^{(\mathbf{x})}$的 j^th 个元素，它们是编码器网络 e_ϕ 在给定输入**x**时的输出；D 是潜向量 z的长度。

L_r用于量化输入**x** 与解码器网络 d_ψ的输出之间的差异，其定义如下:

$$L_\text{r}(\mathbf{x};\phi,\psi) = \sum_{i=1}^{N} x_i \log(\mathbf{x'}_i) + (1-\mathbf{x}_i)\log(1-\mathbf{x'}_i)$$

其中，$\mathbf{x}_i$ 是输入**x**中第 i^th 个单元的值，$\mathbf{x'}_i$ 是解码器输出$\mathbf{x'} = d_\psi(z^{(x)})$中第 i^th 个单元的激活值，其中 $z^{(\mathbf{x})} = \mu^{(\mathbf{x})} + \sigma^{(\mathbf{x})} * \epsilon$，$\epsilon$从 $X \sim \mathcal{N}(0, I)$中采样，N是**x**的长度。

L_c m衡量 c_ω的输出与目标输出 y 之间的差异。当模型进行训练时，在实际任务上，其中 y 是一个硬目标（即独热编码），L_c 定义为:

$$L_\text{c}(\mathbf{x},\mathbf{y};\phi,\psi,\omega) = -\log p(Y=\mathbf{y}|\mathbf{x})$$

其中，p 是网络针对输入**x**输出的条件概率分布。具体而言，p 对应于 c_ω的 softmax 输出。在任务增量学习而非领域增量学习和类别增量学习场景中，

如先前工作 [34]所示， c_ω 的 softmax 输出仅激活表示当前任务或重放任务中类别的输出单元。在对每个重放样本（**x**'、**y**'）进行训练时，其中**y**′ 充当"软"目标，为每个活跃类别提供一个概率， L_{c} 被定义为蒸馏损失 L_{d}:

$$L_{\mathrm{d}}(\mathbf{x}', \mathbf{y}'; \phi, \psi, \omega) = -T^2 \sum_{i=1}^{M} \mathbf{y}'_i \log p^T(Y = i|\mathbf{x}')$$

其中， M 是所考虑的类别数量， T 是温度参数，用于控制模型输出的概率的相对缩放。对于回放，**y**′ 是使用主模型在最近任务训练完成后保存的副本获得的。

A.3 训练细节

在我们的模拟中，split-MNIST任务的批量大小为128，拆分 CIFAR任务的批量大小为256。模型在每个split-MNIST任务中进行2000次训练迭代，在每个拆分 CIFAR任务中进行5000次训练迭代。在split-MNIST模拟中，我们使用学习率为0.001的Adam 优化器（$\beta_1 = 0.9,\ \beta_2 = 0.999$），在拆分 CIFAR模拟中则使用学习率为0.0001。回放期间，我们采用相同的迭代次数和批量大小。

在线训练期间，模型通过最小化 L_{total}进行学习。离线重放期间，生成器（即变分自编码器）对 x' 进行采样，分类器将 x' 标记为 y'。给定作为输入的 x'，模型通过将 y' 作为目标、将 L_c 作为蒸馏损失 L_d，以最小化 L_{total} 为训练目标。
在所有模拟中，模型都会学习一系列分类任务，每个任务都包含数据集（即 MNIST 或 CIFAR-100）中的一部分类别。在拆分 MNIST 中，10 个数字类别被分为 5 个任务，每个任务都是一个二分类任务。相比之下，拆分 CIFAR-100 包含 10 个任务，每个任务都是一个十分类任务。

在类别场景中（但不包括任务场景和领域场景），任务之间，模型会生成重放数据，并将其与任务输入交错呈现。在所有场景中，所有任务呈现完毕后，模型会在离线阶段利用自身副本生成的重放数据进行学习。

A.4 测试方案

每个任务的模型性能通过 8 个批次进行测试，每批次包含 128 个样本（即每个任务共 1024 张图像）。对于使用不同随机种子初始化的模型，无论是在所有任务上进行离线重放之前还是之后，均使用完全相同的测试数据评估其性能。
沿用先前的工作 [35], ，我们在三种情境下评估了模型的性能：任务增量学习、领域增量学习和类别增量学习。三种情境的主要区别在于测试时任务身份的可用程度。以一个 split-MNIST任务为例，模型依序学习五个二分类数字分类任务。在任务增量学习情境下，给定一张输入图像，模型会从包含该图像实际类别的任务所对应的两个类别中进行选择。在领域增量学习情境下，模型会在不知道任务身份的情况下，判断图像属于某个任务的第一个还是第二个类别。最后，在类别增量学习情境下，模型会从目前为止见过的全部 10 个类别中进行选择。